LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-608237

A Best Fit Approach to Estimating Multiple Diffuse Source Terms Using Ambient Air Monitoring Data and an Air Dispersion Model

D. H. MacQueen, N. A. Bertoldo, A. M. Wegrecki

December 14, 2012

Operational Radiation Safety

Disclaimer

LLNL-JRNL-608237

A Best Fit Approach to Estimating Multiple Diffuse Source Terms Using Ambient Air Monitoring Data and an Air Dispersion Model

Donald H. MacQueen, Nicholas A. Bertoldo, and Anthony M. Wegrecki*

This document contains the content, but not the layout, of the version that will be published in the journal Operational Radiation Safety, based on the final proof received from the journal.

correspondence:

Donald H MacQueen

Lawrence Livermore National Laboratory
7000 East Ave. L-627
Livermore, CA 94550

macqueen1@llnl.gov

925-423-1062 (tel)
925-422-2748 (fax)

ABSTRACT

Lawrence Livermore National Laboratory uses CAP88-PC Version 1.0 modeling software to demonstrate compliance with the Code of Federal Regulations Title 40 Part 61 Subpart H (National Emission Standards for Emissions of Radionuclides Other Than Radon From Department of Energy Facilities). Annual air emissions from both well characterized stack sources and difficult to characterize diffuse sources must be assessed. This paper describes a process that uses a mathematical optimization routine to find a set of estimated diffuse source terms that together with the measured stack source terms provides a best fit of modeled air concentrations to measured air concentrations at available sampling locations. The estimated and measured source terms may then be used in subsequent CAP88-PC modeling to estimate dose at the off-site maximally exposed individual. LLNL has found this process to be an effective way to deal with the required assessment of diffuse sources that have otherwise been difficult to assess.

INTRODUCTION

The Lawrence Livermore National Laboratory (LLNL) is a premier research laboratory that is part of the National Nuclear Security Administration (NNSA) within the Department of Energy (DOE). The LLNL Livermore Site is located at the eastern border of the Livermore Valley in northern California, approximately 65 km (40 miles) east-southeast of San Francisco. Title 40 of the Code of Federal Regulations Part 61 Subpart H, National Emission Standards for Hazardous Air Pollutants (RADNESHAPs) requires LLNL to demonstrate yearly that radionuclide emissions to ambient air do not exceed levels that would result in an annual effective dose equivalent of 100 µSv to any member of the public. Compliance is demonstrated using EPA-approved CAP88-PC software to estimate the airborne dispersion of radionuclides from emission sources and resultant radiological dose and risk. The RADNESHAPs regulations apply to all radionuclide emissions other than radon; however, the methods in this paper were developed for tritium emissions, specifically tritiated water vapor (HTO).

In addition, and by Department of Energy order, LLNL operates an environmental monitoring program that includes monitoring tritium in ambient air moisture using silica-gel air samplers and, in some facilities, stack monitoring using glycol bubblers. These networks provide information regarding ambient air concentrations of radionuclides that may contribute to a dose to the public. The ambient air network measures HTO only; whereas stack monitoring measures both HTO and tritiated hydrogen gas (HT).

Potential sources of tritium releases to ambient air at LLNL include both stack sources, such as facilities using tritium in research activities, and diffuse sources, such as materials stored in waste accumulation areas. However, RADNESHAPS Subpart H provides procedures only for stack sources (USDOE 1995). In 1995 the DOE and the US Environmental Protection Agency (EPA) agreed that DOE facilities would also evaluate and report emissions from diffuse sources, but procedures were not specified (USDOE 1995). This paper describes LLNL's current procedure for evaluating HTO emissions from diffuse sources. The procedure and its results are described in LLNL's annual RADHESHAPS reports.

LLNL uses the EPA-approved CAP88-PC Version 1.0 software (USEPA 1992) (hereafter referred to as CAP88-PC) to estimate annual doses to members of the public and demonstrate RADNESHAPs compliance. LLNL operations generally include both stack and diffuse sources. The stack monitoring instruments measure the average air concentration and the total airflow in each monitored stack every week. The total activity released is calculated from these data. Diffuse sources, on the other hand, might consist of for example storage units containing tritium-contaminated items located in a waste accumulation area. Although the quantity of contaminated items might be known, the activity released to the environment is not well characterized, and cannot be readily measured. Thus, operational information can provide only a rough estimate of the source term input necessary for CAP88-PC.

This paper describes a process to simultaneously estimate multiple diffuse source terms that, when combined with the measured stack source terms, minimizes the root mean square (*rms*) of the differences between modeled and measured annual average HTO air concentrations at ambient air sampling locations. Because CAP88-PC dose estimates are

derived from its concentration estimates, this process works with CAP88-PC concentration output (later versions of CAP88-PC have updated dose but not concentration calculations, and thus could be used instead of version 1). The essence of the method is to hold the stack source terms fixed and vary the diffuse source terms until the best fit of modeled air concentrations to measured air concentrations is obtained. Although the idea is conceptually simple, the details are relatively complex. The process is implemented using scripts and functions written in the R programming language (R Core Team 2012). The search for best-fit estimates of diffuse source terms is optimized by mathematical routines from the PORT library (Fox 1984) available within R.

The process is detailed in the remainder of this paper and can be summarized as follows:

1. Run CAP88-PC for each stack using its known source term and for each diffuse source using a unit source term to obtain source-specific modeled concentration values at the CAP88-PC radial grid sector locations.
2. Using interpolation, map the CAP88-PC radial grid results to a rectangular receptor grid and then obtain source-specific modeled concentration values at the ambient air monitoring station locations.
3. Input the results from Step 2 and the measured ambient air concentrations at each monitoring location into the software (R scripts).
4. Keeping the stack source concentration contributions fixed, use the software to scale each diffuse source contribution to obtain a best fit of modeled to measured air concentrations.

5. Since the diffuse source models in Step 1 used a unit source, the scale factors determined in Step 4 are in effect the estimated diffuse source terms.
6. Re-run CAP88-PC using the measured stack source terms and the estimated diffuse source terms to estimate dose.

METHODS

CAP88-PC grid geometry

CAP88-PC allows the user to input a set of up to 20 distances from a source location. It then creates a radial grid of points centered on (0,0) at those distances in each of 16 compass directions. Fig. 1 illustrates a typical radial grid centered on (0,0).

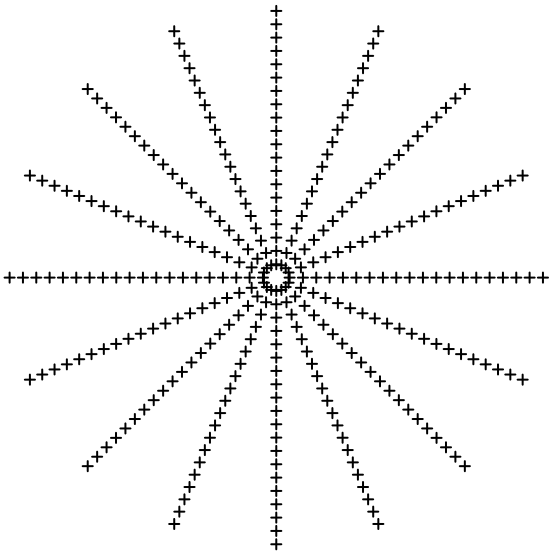

Fig. 1. Example CAP88-PC radial grid.

Interpolating and adding multiple CAP88-PC runs

When combining output from two or more CAP88-PC runs representing different sources, each must have an appropriate center so that the relative positions of the sources and radial grid points are correct. In general, the points of two or more CAP88-PC radial grids will not coincide. Similarly, points on radial grids will not coincide with ambient air sampling

locations. This is illustrated in Fig. 2, which shows radial grids from two LLNL sources along with a number of ambient air sampling locations relative to the LLNL site boundary.

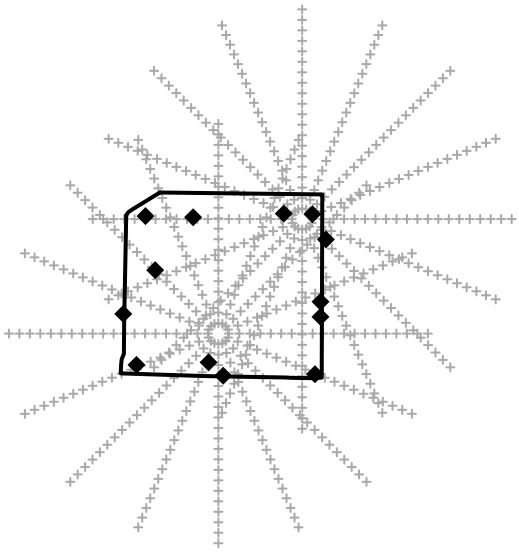

Fig. 2. Two CAP88-PC radial grids with ambient air sampling locations, relative to the LLNL site boundary.

Because sampling locations and multiple radial grid locations do not coincide, model concentrations at sampling locations are calculated by interpolation and addition. First, a rectangular grid of equally spaced receptor locations is superimposed on the entire system. Next, concentrations are interpolated separately from each CAP88-PC radial grid onto the rectangular grid and added. Finally, concentrations at each sampling location are interpolated from the rectangular grid. Interpolation is in log units because model concentrations range over several orders of magnitude and are not linear as a function of distance from a source.

Interpolation is possible only where all radial grids overlap. The examples in Figs. 1 and 2 use 100 meter spacing along the radials, from 100 m to 2,000 m. The overlapping area is smaller than either of the radial grids alone and does not cover the entire site (the grid centered in the northeast corner does not quite reach the southwest corner of the site). One

could increase the grid spacing, but the greater the grid spacing the coarser the interpolation. We have chosen to run CAP88-PC twice for each source: once with radial distances from 100 m to 2,000 m, and again with distances from 2,100 m to 4,000 m. The output from the two runs is then assembled as if it were a single run with distances from 100 m to 4,000 m (this is necessary only because CAP88-PC is limited to 20 distances in a single run). Using such assembled runs ensures that the intersection of the radial grids covers the entire site and well beyond, while maintaining relatively short distances between points along the radii.

The Fitting Process

Input to the fitting process consists of the following:

- CAP88-PC concentration output from the sources of interest. These take the form of (*x,y,z*) triples, where (*x,y*) are the coordinates of the radial grid points, and (*z*) is the model concentration at (*x,y*).
    - Stack sources models use their respective known source terms.
    - Diffuse source models use unit source terms.
- Ambient air sampling locations, in the form of (*x,y*) coordinate pairs.
- Measured concentrations at the sampling locations.
- An equally-spaced grid represented by (*x,y*) pairs that encompasses all receptor locations of interest (a 100x100 grid would have 10,000 points).

All inputs must represent the same time period. Normally, this is one year, in which case the measured concentrations are annual averages, the CAP88-PC wind file is derived from annual meteorology, and CAP88-PC model concentrations represent annual averages.

The user supplies the optimization software with initial guesses for the diffuse source terms. The software performs the interpolation described above. The interpolated model concentrations are compared with measured concentrations. Goodness of fit is measured by the root-mean-square differences between modeled and measured concentrations:

$$rms = \sqrt{\frac{\sum_{i=1}^{n} (\hat{c}_i - c_i)^2}{n}} \qquad \text{Eq. 1}$$

where

$\hat{c}_i$ = model concentration at location i

c_i = measured concentration at location i.

The software then internally tries new choices for the diffuse source terms, performs the interpolations, and calculates a new *rms*. It iteratively repeats this process using a quasi-Newton algorithm[i] until it converges on a smallest value for the *rms*. The source term values that produce the smallest *rms* are returned as the best-fit estimates of the diffuse source terms.

Note that the CAP88-PC diffuse source model concentrations change with every choice of diffuse source terms and therefore must be recalculated repeatedly during the iterations. This is easy to do because the diffuse source models are run with a unit source and concentration scales linearly with the source term. At each iteration the "current choice" for the diffuse source term is actually a scaling factor, and the input unit source model concentration (*z*) values are multiplied by the scaling factor before interpolating. Thus it is not necessary to actually re-run the CAP88-PC software at each iteration. This greatly simplifies the computer programming.

EXAMPLE

Source and ambient air monitoring locations

The process is illustrated with three stack sources and four diffuse sources (Fig. 3). The ambient air tritium monitoring network includes thirteen locations on or adjacent to the site (ARAC, B295, CAFE, COW, CPET, CRED, DWTF, MESQ, MET, POOL, SALV, SECO, and VIS), also shown in Fig. 3.

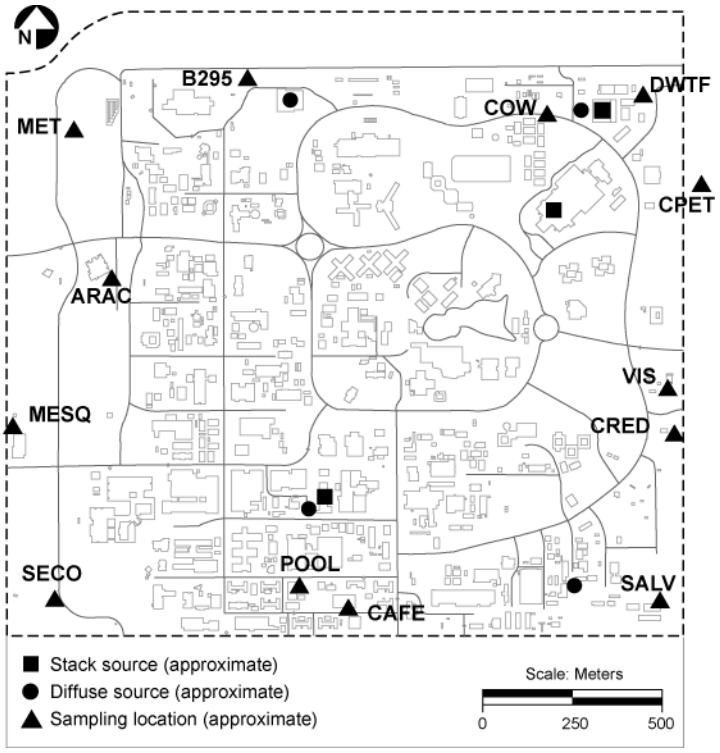

Fig. 3. LLNL stack sources, diffuse sources, and ambient air tritium sampling locations.

On an annual basis winds at LLNL are predominantly from the southwest (Fig. 4), hence dispersion is expected to be predominantly toward the northeast.

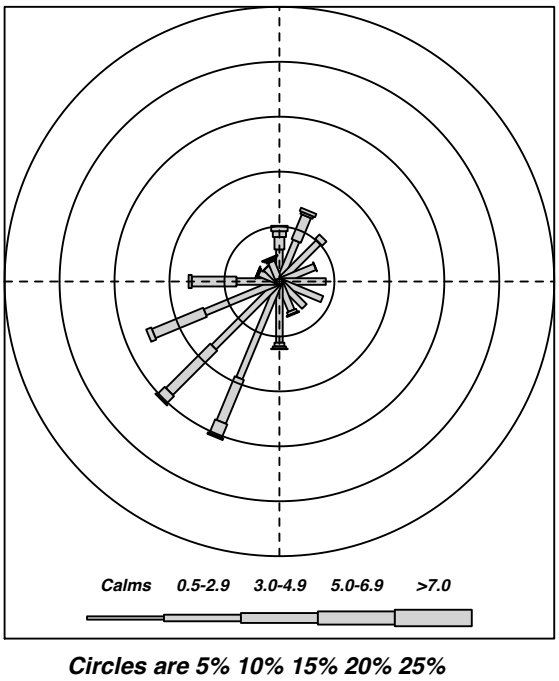

Fig. 4. Example LLNL windrose showing wind speeds (m s^{-1}) and directions from which winds blow.

Interpolation and addition using two sources

Figs. 5a and 5b show the CAP88-PC concentration field for one stack source and one diffuse source respectively, each interpolated individually from their respective radial grids onto a rectangular grid. The diffuse source model used a unit source term. Both images use the same concentration scale, from about −7 to −1 in $\log_{10}$ concentration units.

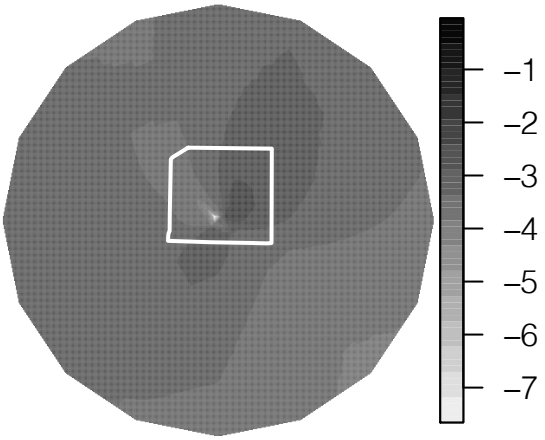

Fig. 5a. CAP88-PC concentration field for one example stack source ($\log_{10}$ Bq m^{-3}; same scale as Fig. 5b).

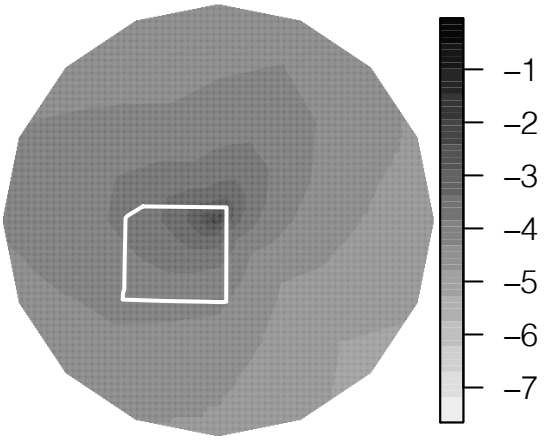

Fig. 5b. CAP88-PC concentration field for one example diffuse source ($\log_{10}$ Bq m^{-3}; same scale as Fig. 5a).

As described above, a key step in the fitting process is to take the concentration output from all of the sources on their respective radial grids, interpolate each to a common regular grid, and add. The result of this step using the two sources of Figs. 5a and 5b is shown in Fig. 6. The range of concentrations in Fig. 6 has been reduced to show more detail. The concentration field in Fig. 6 shows elements from both sources. Note also the oblong shape of overlap between the two radial grids. Combining only two of seven sources, as in Fig. 6, is not an essential step in the fitting process, but does provide a useful reality check on whether the

interpolate and add process is working correctly.

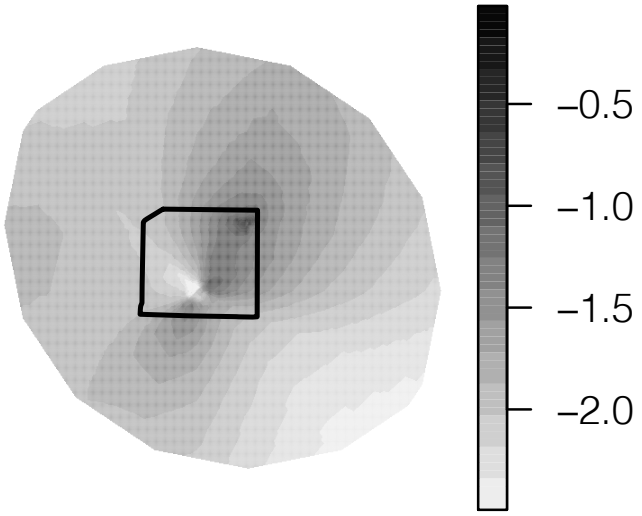

Fig. 6. Example of one stack and one diffuse sources interpolated to a common grid and added (log_{10} Bq m^{-3}).

Best fit using seven sources

Fig. 7 illustrates the result of the same process, but using all seven sources and with the diffuse source terms adjusted to obtain the best fit of modeled to measured concentrations. Note that the spatial pattern of concentrations is now more complex, as should be expected when multiple sources are present.

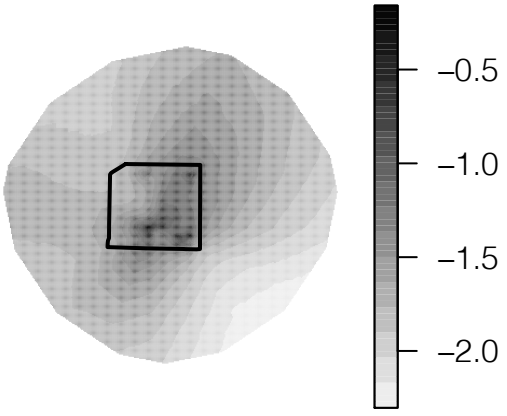

Fig. 7. Best fit result using all seven example sources (log_{10} Bq m^{-3}).

Estimated diffuse source terms in this example range from 4.1 to 35.1 Bq. The fit of modeled to measured annual average ambient air concentrations at the sampler locations is shown in Fig. 8.

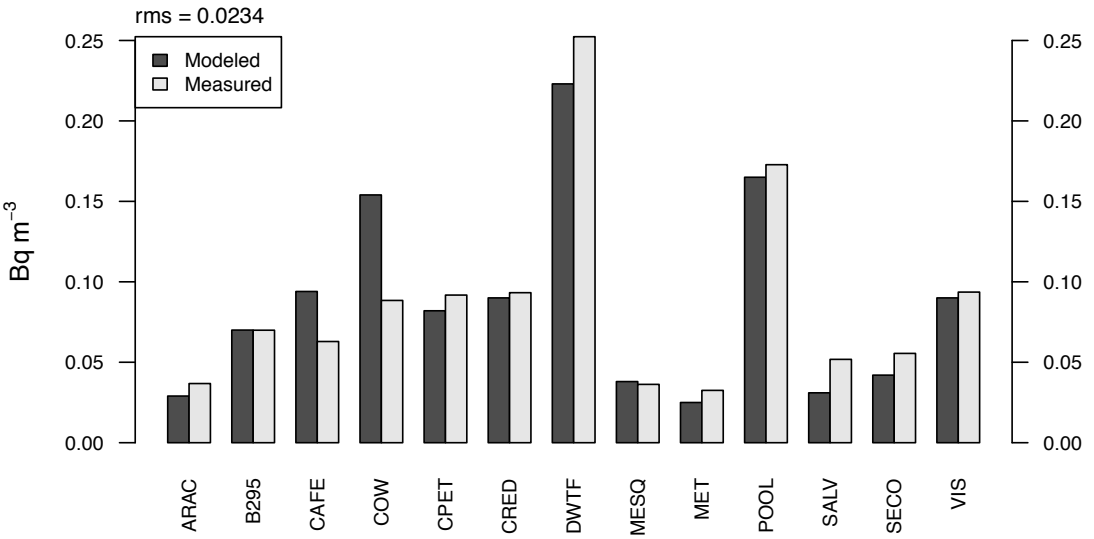

Fig. 8. Comparison of measured and best-fit model annual average concentrations of tritiated water vapor (HTO) in ambient air (Bq m^{-3}). *rms* = 0.0234 (Eq. 1).

The modeled concentration is within a factor of two of the measured concentration at all locations, and within 20% at seven of thirteen.

Further evaluation of the best-fit model.

Fig. 9 compares the modeled and measured concentrations when only the stack sources are used. This can be used to evaluate the contribution of the diffuse sources.

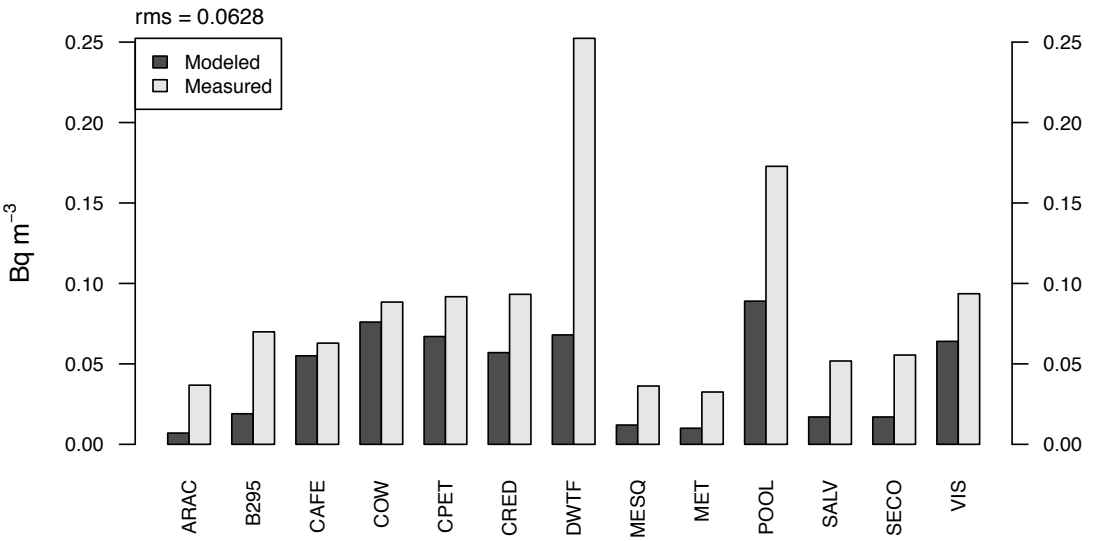

Fig. 9. Comparison of measured and modeled annual average concentrations using stack sources only. *rms* = 0.0628 (Eq. 1).

The "stack sources only" *rms,* 0.0628, is larger than the best-fit *rms,* 0.0234. The modeled stack sources alone underestimate the measured concentrations at every sampling location, and drastically so at DWTF and POOL. Comparing Fig. 9 with Fig. 8 suggests that stack sources alone do not account for the measured concentrations, and that stack and diffuse sources together can better account for the measured concentrations. This makes sense, since stack sources are not the only sources at LLNL.

Another way to show the model fit is a scatterplot of modeled vs. measured results. Figs. 10 and 11 illustrate this for the best-fit model and "stack sources only" model respectively. A line with slope = 1 and intercept = 0 is superimposed on each scatterplot.

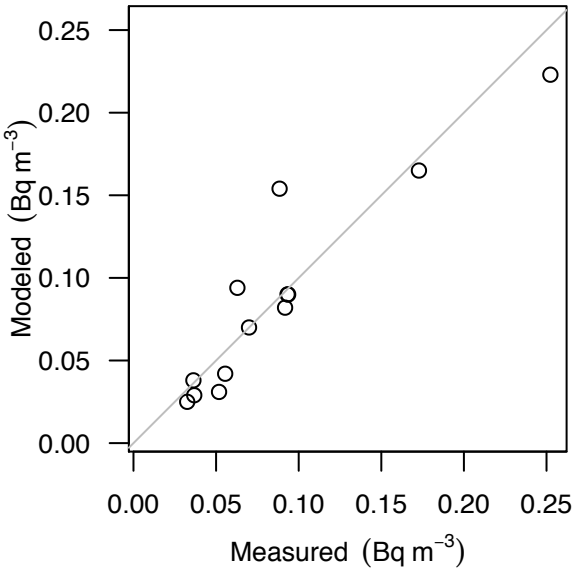

Fig. 10. Scatterplot of best fit results. r^2= 0.85.

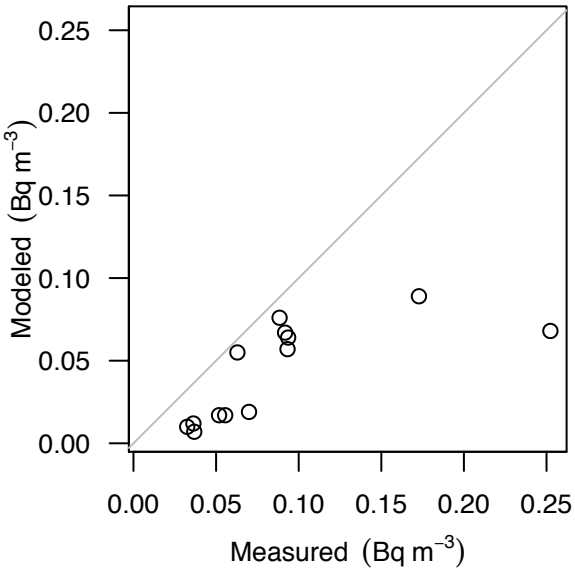

Fig. 11. Scatterplot of stack sources only. r^2= 0.5.

Like Fig. 9, Fig. 11 shows that the "stack sources only" model under-predicts at every location.

DISCUSSION

Considerations for correct implementation

As noted above, it is important that all of the inputs to the fitting process represent the same time period. Normally, the time period is one calendar year, as this is the basis for RADNESHAPs compliance. However, operational changes, such as changes of sampling locations, the introduction of a new source, or the removal of an existing source, can introduce different time periods. It is possible, but cumbersome, to run this process separately for different portions of a year.

In addition, it is important that the inputs represent the appropriate spatial resolution. The minimum distance supplied to CAP88-PC should be less than the minimum distance from any source to any sampling location. Otherwise, the interpolation of "close-in" sampling locations is based entirely on concentration output at distances further from the source, and will tend to underestimate the concentration at that location. Other model inputs that affect concentration, such as release height, source area, meteorology, etc., must be appropriate. Operational information should guide model inputs as much as possible.

Although model bias is not addressed in this paper, it is important to consider its potential effect. For example, if a model consistently overestimates concentrations then the fitting process will compensate by underestimating source terms.

CAP88-PC calculates concentrations in air that result from a specific modeled source. However, there could be other factors contributing to ambient air concentrations. For example, if this method were applied to a particulate radionuclide, resuspension from previous deposition could contribute to ambient air concentrations. In the case of tritium, conversion of

HT to HTO after release but before sampling could contribute to air sampling results, but would not be represented by the modeling results. In order to use the methods of this paper for RADNESHAPs compliance, any such contributions should be small relative to direct contributions from modeled sources.

In the example above, only HTO emissions were used for stack sources because the ambient air samplers measure HTO only and it is difficult to determine the fraction of HT converted to HTO in the environment. A comprehensive review of the latter issue is beyond the scope of this paper, but a few key points should be mentioned. There are three conversion processes that have been considered in regard to using this method for HTO.[ii] The first is the self-oxidation of tritium; it is well accepted that this process requires a significant amount of available radioactivity and has a very slow rate of reaction (Casaletto et al. 1962). The second is the process of photochemical oxidation, which converts some fraction of HT to HTO. This conversion also occurs slowly, and has been estimated to be less than 1% in the first 24 hours following a release (Burger 1976). The third conversion process occurs when HT diffuses into soil and is oxidized via a bacterial enzymatic (hydrogenase) reaction. In closed soil system studies by McFarlane et al. (1978), the conversion to HTO in soil proceeded to completion within 92 hours. Because of the much longer time scales of the first two processes, their contributions to local ambient HTO measurements represent an insignificant contribution. The time scale measured by MacFarlane and the fact that ambient air tritium sampling is conducted on a bi-weekly basis suggest that the ambient air sampling captures the fraction of HT converted to HTO via soil bacteria. The estimation of the fraction of HT converted to HTO is the subject of future work, however, a preliminary investigation based on deposition rates and

conservative conversion assumptions suggests that ambient HTO air concentrations resulting from HT converted to HTO in soil are very low relative to HTO concentrations from the modeled sources.

Interpretation

The role of CAP88-PC in RADNESHAPs compliance is to estimate doses. CAP88-PC dose estimates are derived from its concentration estimates. It follows, therefore, that improvements to its concentration estimates will improve its dose estimates.

As noted above, modeled concentrations from the LLNL stack sources alone underestimate the measured concentrations. Assuming that the stack source model is not biased low[iii], it follows that contributions from other sources are required to account for the measured concentrations. Furthermore, because LLNL is the only tritium source in the area and ambient air background tritium levels are insignificant, the only other sources available to account for the measured concentrations are LLNL diffuse sources. Thus, modeling both the measured emissions from the stack sources combined with the estimated emissions from the diffuse sources with release heights significantly closer to ground level results in a more conservative and likely more realistic dose estimate at the location of a hypothetical maximally exposed individual.

Strictly speaking, the improvement of concentration estimates is demonstrated only at the ambient air sampling locations. Improvement at other locations within the model domain is inferred from a belief that the model performs well in terms of spatial patterns. At LLNL, modeled peak concentrations are typically on-site or near the fence line. Therefore, ambient air

sampling locations on or near the fence line (as is the case at LLNL) are appropriate for use in this process.

Verification

The R script calculations have been reproduced using spreadsheets and a series of source-specific CAP88-PC runs. Source-specific runs are necessary because the distances and directions from the sources to the samplers are different for each source. CAP88-PC lets the user specify exact distances to the samplers, but not exact directions (directions are fixed at N, NNW, NW, WNW, etc.), so it is not possible for CAP88-PC to model concentrations at each exact sampler location. Instead, the modeled sampler concentration from a source-specific model run will be that of the sector containing the sampler. The optimization method, in contrast, interpolates the sampler concentrations from the nearby sector centers. Allowing for the difference in interpolation approach, the optimization results and the results using source-specific runs and spreadsheets agree.

Additional uses of the method

After a best fit is obtained, the resulting regular grid of concentration estimates is available for other uses. One such use is to interpolate concentration profiles along paths of interest, such as the site boundary. Fig. 12 shows such a profile, starting in the northeast corner of the LLNL site and proceeding clockwise around the site. The dip at about 1,800 m corresponds to the southeast corner; the maximum at about 3,000 m is along the southern border, and the extended trough from about 3,500 m to 6,000 m represents the west and northwest boundaries. These patterns are consistent with prevailing wind directions (Fig. 4).

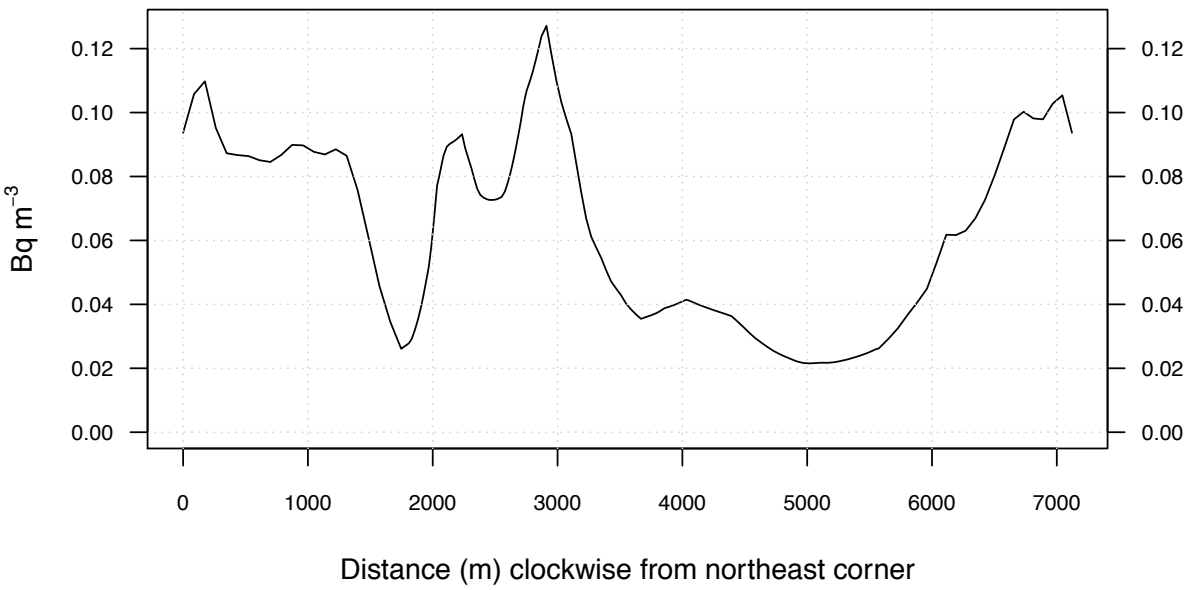

Fig. 12. Modeled concentration profile along the LLNL fence line.

Finally, it should be noted that the numerical optimization process itself is entirely concentration-based and therefore does not depend on which radionuclide is being modeled. This method can be applied to any radionuclide.

Future work

It might be useful to include HTO emissions from soil in the best-fit process by treating large soil areas as additional diffuse sources. If doing so gave a significantly better fit to the measured data, this would suggest that perhaps the conversion of HT to HTO in soil deserves further study.

SUMMARY

This paper presents a method for simultaneously estimating source terms for multiple diffuse sources by optimizing the fit of model concentrations to measured concentrations at multiple sampling locations. The method has worked well at LLNL for the past several years, and its estimates have been reasonable based on process knowledge.[iv]

ACKNOWLEDGMENTS

The authors would like to thank LLNL staff members Gary Bear, Steve Hall, Terrance Poole, and Kent Wilson for their work designing, operating, and maintaining the environmental sampling and meteorological monitoring equipment that provide the information necessary for this work. The authors also thank the reviewers for their helpful and thought-provoking comments. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory in part under Contract W-7405-Eng-48 and in part under Contract DE-AC52-07NA27344.

REFERENCES

Burger LL. The conversion of tritiated hydrogen to water in the atmosphere. Richland, WA: Pacific Northwest National Laboratory; BNWL-2113, UC-20; 1976.

Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. SIAM J. Sci. Comput. 16:1190-1208; 1995.

Casaletto GJ, Gevantman LH, Nash JB. The self-radiation oxidation of tritium in oxygen and air. Naval Radiological Defense Laboratory Report USNRDL-TR-565; San Francisco, CA; May 8, 1962.

Fox, PA, ed. The PORT mathematical subroutine library. AT&T Bell Telephone Laboratories, Inc.; 1984. Available at http://www.netlib.org/port. Murray Hill, NJ. Accessed 9 November 2012.

Hill RL, Johnson JR. Metabolism and dosimetry of tritium. Health Phys 65:628-647; 1993.

McFarlane JC, Rogers RD, and Bradley DV Jr. Environmental tritium oxidation in surface soil. Environ. Sci. Technol. 12:590-593; 1978.

Nocedal J, Wright SJ. Numerical optimization. New York: Springer; 1999.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL http://www.R-project.org/; 2012. Accessed 6 March 2013.

U.S. Department of Energy, U.S. Environmental Protection Agency. Memorandum of Understanding between the U.S. Environmental Protection Agency and the U.S. Department of Energy concerning The Clean Air Act Emission Standards for Radionuclides 40 CFR Part 61 Including Subparts H, I, Q & T; May 16, 1995. Available at http://yosemite.epa.gov/rpd/neshaps_pid.nsf/0/959ECF89A5B7134486256F1C00677448?Opendocument. Accessed 9 November 2012.

U.S. Environmental Protection Agency. User's guide for CAP88-PC Version 1. EPA 402-B-92-001; 1992. Available at http://www.epa.gov/rpdweb00/docs/cap88/402-b-92-001.pdf. Accessed 20 December 2012.

U.S. Environmental Protection Agency. About CAP-88 PC [online]; April 26, 2012. Available at: http://www.epa.gov/radiation/assessment/CAP88/aboutcap88.html#validation. Accessed 6 March 2013.

* Lawrence Livermore National Security, LLC, 7000 East Ave., Livermore, CA 94550

[i] The quasi-Newton method is implemented in the R function "optim". The method uses *rms* values (Eq.1) and gradients to develop a multi-dimensional view of the value to be optimized (in this case, *rms* as a function of various possible values for the diffuse source terms). Per the R manual page for the optim function, refer to Byrd (1995) and Nocedal (1999) for more detail.

[ii] Although tritiated molecular combinations occur as gas, water, methane (Hill 1993) and other compounds, this discussion will center on the conversion of HT to HTO because HT is a measured stack emission and HTO is the modeled radionuclide.

[iii] According to the US EPA, "EPA has made comparisons between the predictions of annual-average ground-level concentration to actual environmental measurements, and found very good agreement" (USEPA 2012). This suggests that CAP88-PC does not systematically underestimate air concentrations.

[iv] R scripts and functions are available from the corresponding author.